

Network Analysis of Genome-Wide Association Studies for Chronic Obstructive Pulmonary Disease in the Context of Biological Pathways

Citation for published version (APA):

Mount, S., Cirillo, E., Stewart, K., Coort, S., Evelo, C. T., Wesselius, A., Zeegers, M. P., & Schols, A. M. W. J. (2019). Network Analysis of Genome-Wide Association Studies for Chronic Obstructive Pulmonary Disease in the Context of Biological Pathways. *American Journal of Respiratory and Critical Care Medicine*, 200(11), 1439-1441. <https://doi.org/10.1164/rccm.201904-0902LE>

Document status and date:

Published: 01/12/2019

DOI:

[10.1164/rccm.201904-0902LE](https://doi.org/10.1164/rccm.201904-0902LE)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

ORCID IDs: 0000-0002-3677-1996 (R.S.J.E.); 0000-0001-9686-5698 (F.P.).

*These authors contributed equally to this work.

†F.D.M. is Deputy Editor of *AJRCCM*. His participation complies with American Thoracic Society requirements for recusal from review and decisions for authored works.

§Corresponding author (e-mail: fpolverino@copdnet.org).

References

1. Soriano JB, Polverino F, Cosio BG. What is early COPD and why is it important? *Eur Respir J* 2018;52:1801448.
2. Woodruff PG, Barr RG, Bleecker E, Christenson SA, Couper D, Curtis JL, et al.; SPIROMICS Research Group. Clinical significance of symptoms in smokers with preserved pulmonary function. *N Engl J Med* 2016; 374:1811–1821.
3. Petersen H, Vazquez Guillemet R, Meek P, Sood A, Tesfaigzi Y. Early endotyping: a chance for intervention in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2018;59:13–17.
4. Hogg JC, Chu F, Utokaparch S, Woods R, Elliott WM, Buzatu L, et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004;350:2645–2653.
5. Núñez B, Sauleda J, Antó JM, Julià MR, Orozco M, Monsó E, et al.; PAC-COPD Investigators. Anti-tissue antibodies are related to lung function in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011;183:1025–1031.
6. Lee SH, Goswami S, Grudo A, Song LZ, Bandi V, Goodnight-White S, et al. Anti-elastin autoimmunity in tobacco smoking-induced emphysema. *Nat Med* 2007;13:567–569.
7. Polverino F, Cosio BG, Pons J, Lacho-Contreras M, Tejera P, Iglesias A, et al. B cell-activating factor: an orchestrator of lymphoid follicles in severe chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2015;192:695–705.
8. Faner R, Cruz T, Casserras T, López-Giraldo A, Noell G, Coca I, et al. Network analysis of lung transcriptomics reveals a distinct B-cell signature in emphysema. *Am J Respir Crit Care Med* 2016;193: 1242–1253.
9. Müller NL, Staples CA, Miller RR, Abboud RT. “Density mask”: an objective method to quantitate emphysema using computed tomography. *Chest* 1988;94:782–787.
10. Cosio MG, Saetta M, Agusti A. Immunologic aspects of chronic obstructive pulmonary disease. *N Engl J Med* 2009;360: 2445–2454.

Copyright © 2019 by the American Thoracic Society



Network Analysis of Genome-Wide Association Studies for Chronic Obstructive Pulmonary Disease in the Context of Biological Pathways

To the Editor:

Chronic obstructive pulmonary disease (COPD) is a common respiratory disease projected to be the third leading cause of death by 2020 (1). The main risk factor is tobacco smoking, but other

environmental exposures may also contribute (1). Furthermore, host factors including genetic abnormalities, abnormal lung development, and accelerated aging increase susceptibility to COPD (1). However, the causal mechanisms remain poorly understood (2).

As a result of genome-wide association studies, many interesting genetic variations, including SNPs, have been discovered. However, the interpretation of these large amounts of data within the context of biological systems, disease processes, and unknown gene functions remains difficult. Considering genes in a biological context may aid in the elucidation of SNP function. Network analysis provides a way of deciphering the biological relationships among SNPs, genes, and pathways by providing a framework that allows for the integration, analysis, and display of these complex data (3).

We used data from a recent meta-analysis to identify and extract all genetic variants published in pooled and meta-analysis studies related COPD risk (Prospero CRD4201705; May 2018). We extracted the 181 significant genetic variants (regardless of linkage disequilibrium) mapped to 99 genes that included 176 SNPs with reference SNP cluster identifier (rs) and other variants such as multiple SNP combinations, insertions and deletions, or length polymorphisms.

Genes and variants were represented in a SNP–gene network using Cytoscape version 3.6. Second, the genes were used to retrieve the biological pathways from WikiPathways Human curated collection (10 July 2018). Genes present in one or more pathways were displayed in a Cytoscape gene–pathway network. The SNP–gene and gene–pathway networks were then consolidated by merging them. This yielded a SNP–gene–pathway network that was used as a basic reference for the biological interpretation of the connected elements. Finally, genes were classified according to their function and potential effect, using the variant effect predictor analysis in Ensembl (4).

Our analysis produced four different visualizations. In Table 1, an overview of the main characteristics of the networks is reported. In each network title, the digital object identifier to the Network Data Exchange visualization is provided and the main features of the networks and nodes codes are reported, all of which are fully downloadable and interactive.

The networks consist of 181 variant nodes, 99 gene nodes, and 315 pathway nodes, and 735 connections between them. Of the original set of 99 genes, 74 genes are present in pathways from the curated WikiPathways collection. The basic version, Gene–pathway network, highlights the three elements: SNPs, genes, and pathways in different colors.

The Functional gene map visualization presents functional classes in the network. Here we show 13 nonoverlapping functional classes: Addiction, Cellular interaction, Cellular metabolism, Cellular structure, Detoxification, Development, Homeostasis organismal, Inflammation, Lung function, Metabolism organismal, Regulation, Tissue remodeling, and Unknown. Interestingly, some of the gene functional classes are dispersed, whereas in others all are connected. Cellular metabolism (forest green) shows dispersion: 15 genes are not connected in the major central network, and 7 of the 15 do not present any pathway connections. Comparatively, all five genes related to Detoxification cluster in a specific area (refer to online visualization, pink-nodes). Similarly, all 15 genes involved in Inflammation are intensely connected to genes and other pathways and are grouped in the

Author Contributions: S.M., E.C., K.S., S.C., C.T.E., A.W., M.P.Z., and A.M.W.J.S. provided conception and design; S.M., E.C., and K.S. provided analysis and interpretation; S.M. and E.C. drafted the manuscript for important intellectual content; S.M., E.C., K.S., S.C., C.T.E., A.W., M.P.Z., and A.M.W.J.S. provided revision; and S.M., E.C., K.S., S.C., C.T.E., A.W., M.P.Z., and A.M.W.J.S. provided final approval.

Originally Published in Press as DOI: 10.1164/rccm.201904-0902LE on July 25, 2019

Table 1. Overview of the Four SNP–Gene–Pathway Networks

Network Title	Weblink	Features	Node Color
Gene-pathway network	http://doi.org/10.18119/N9SW2N	This is the result of the merged “SNP-gene” and “gene-pathway” networks. This is the basic set for the other two network visualizations.	SNP = green; gene = blue; pathway = red
Functional-gene-map	http://doi.org/10.18119/N9JC76	Genes belonging to a specific functional class listed and color coded.	SNP = gray triangle; gene = squared colored node according to the description; pathway = gray circle
Variant-effect network	http://doi.org/10.18119/N9P301	SNPs with different type of gene impact are highlighted.	SNP low = green; SNP moderate = blue; SNP modifier = orange; gene = gray triangle; pathway = gray circle

Clicking the three weblinks presented in Table 1 will bring you to the interactive open access network visualizations. The Gene-pathway network shows the result of the merged “SNP-gene” and “gene-pathway” networks. The second visualization, Functional-gene-map, is the functional gene map and highlights the groupings of the different functional classifications of the genes. Finally, the Variant-effect network visualization shows the predicted effect of the gene mutations. The color coding system can be found in the tables and is further explained on the web-based visualization by clicking on the nodes or lines. Available information varies according to the visualization but includes node ID, name, matching attribute, functional network area classification, GeneCard external link, synthetic function, general classification, label, Polyphen, Sorting Intolerant from Tolerant, mutation type, consequence, impact, smoking status, and interaction.

central area of the major network (magenta). This network visualization thus gives an indication of the biological process in which the 26 unlinked genes are involved.

We show that most genes within a functional class group together, indicating that they share common pathways, whereas other genes tend to be more dispersed, perhaps dictated by

pleiotropic gene roles. The distinctions have been made in the functional classes based on the (probable) function of the genes. These classifications are subjective but help to guide the important functions of the gene/protein product in the context of COPD.

The Variant-effect network visualization shows the potential effect of the variants on the gene sequence. The resulting network

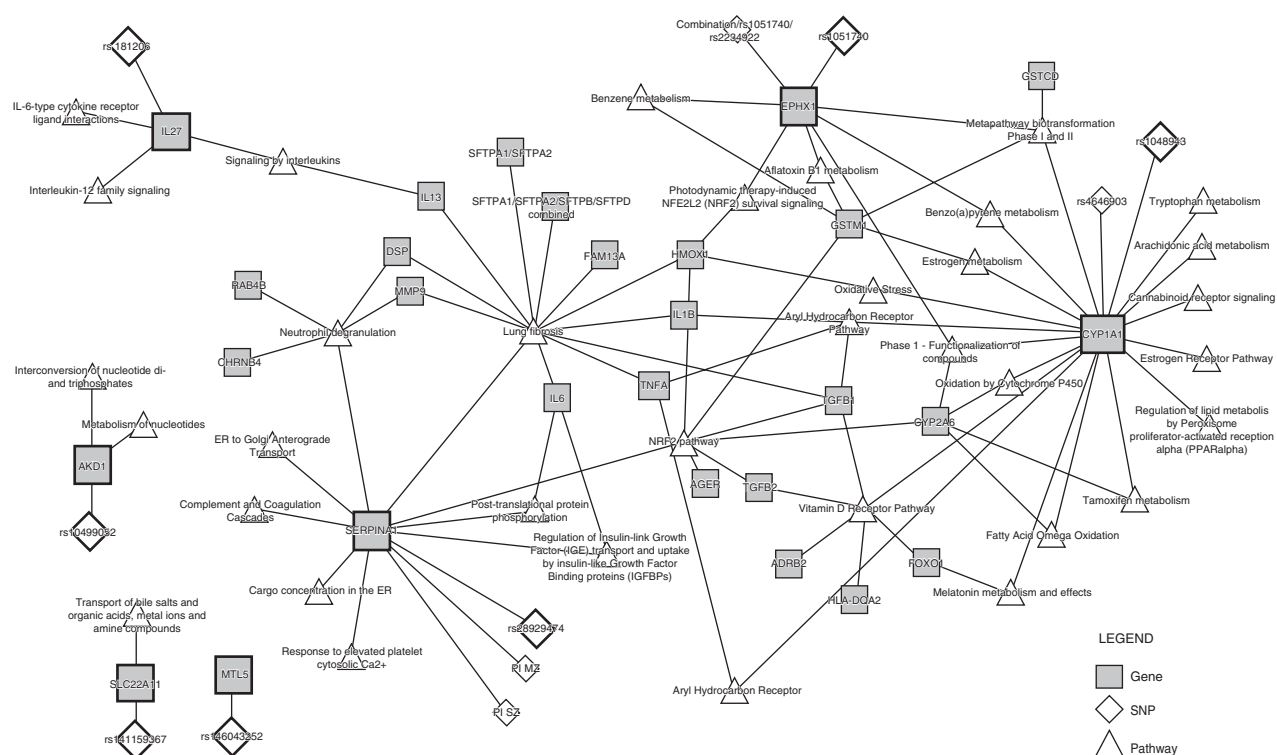


Figure 1. A SNP–gene–pathway subnetwork is presented, highlighting only the connections related to the seven genes represented with bold borders (*AK9*, *SERPINA1*, *IL27*, *CYP1A1*, *EPHX1*, *SLC22A11*, and *TESMIM*) carrying the deleterious missense SNPs.

contains 149 modifiers (or noncoding variants) and 19 moderate (or missense) and two low-impact (or synonymous) SNPs. Polyphen (5) and Sorting Intolerant From Tolerant (6) prediction scores, presented in the network table as attributes, were also consulted to elucidate the deleteriousness of the 19 missense variants. The Polyphen resulted in five probably damaging (rs141159367, rs1051740, rs10499052, rs146043252, and rs28929474) and two possibly damaging (rs1048943 and rs181206) SNPs. Sorting Intolerant from Tolerant indicated that only three (rs1051740, rs10499052, and rs146043252) of the seven SNPs identified by Polyphen are predicted to be deleterious. Because of this discrepancy, we considered the more extensive list. In Figure 1, a SNP–gene–pathway subnetwork is presented, highlighting only the connections related to the seven genes carrying the deleterious missense SNPs.

The variant effect predictor analysis showed that the vast majority of the SNPs associated with COPD are modifiers. Nineteen of the 181 variants were missense SNPs mutations. Of these, seven SNPs (rs10499052, rs28929474, rs181206, rs1048943, rs1051740, rs141159367, and rs146043252) showed alterations predicted to be deleterious in the associated proteins for SLC22A11 (solute carrier family 22 member 11), AK9 (adenylate kinase 9), SERPINA1 (serpin family A member 1), IL27, CYP1A1 (cytochrome P450 family 1 subfamily A member), EPHX1 (epoxide hydrolase 1), and TESMIN (testis-expressed metallothionein-like protein), respectively. Figure 1 displays the interactions of those seven genes with the missense SNPs and pathways. Interestingly, all but two of the deleterious alterations, located in AK9 and TESMIN, are in genes that are either directly or indirectly involved in inflammatory pathways. The AK9 gene mutation is involved in cellular metabolic processes and in extrapulmonary tissues (7), whereas TESMIN is involved in heavy metal ion binding and sequestering.

IL27 and SERPINA1 encode proteins directly involved in inflammation. The leucine to proline substitution caused by rs181206 (IL27) was predicted to be possibly damaging by Polyphen, indicating a strong change in protein structure. Proline is known to have an exceptional conformational rigidity, often causing structural changes. SERPINA1 mutations account for around 2% of all COPD cases (8); however, the PiMZ variant associated with this gene was only observed in crude estimates and disappeared after adjusting for smoking (9). Two deleterious SNPs, rs1048943 and rs1051740, associated with CYP1A1 and EPHX1, are involved in detoxification pathways and, under some circumstances, may be directly linked, as in benzo(a)pyrene-metabolism (WikiPathways identifier: WP696). ■

Author disclosures are available with the text of this letter at www.atsjournals.org.

Sarah Mount, M.Sc.*
Elisa Cirillo, Ph.D.*
Kelly Stewart, Ph.D.
Susan Coort, Ph.D.
Chris T. Evelo, Ph.D.
Anke Wesselius, Ph.D.
Maurice P. Zeegers, Ph.D.
Annemie M. W. J. Schols, Ph.D.*
Maastricht University Medical Centre Maastricht
Maastricht, the Netherlands

ORCID IDs: 0000-0002-3197-0574 (S.M.); 0000-0002-0241-7833 (E.C.); 0000-0002-4451-6012 (K.S.); 0000-0003-1224-9690 (S.C.); 0000-0002-5301-3142 (C.T.E.); 0000-0003-4474-9665 (A.W.); 0000-0002-2387-083X (M.P.Z.); 0000-0001-9878-0428 (A.M.W.J.S.).

*These authors contributed equally/shared first authorship.

†Corresponding author (e-mail: a.schols@maastrichtuniversity.nl).

References

1. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. 2019 [accessed 2019 Apr 12]. Available from: <https://goldcopd.org>.
2. Celli BR, Decramer M, Wedzicha JA, Wilson KC, Agustí A, Criner GJ, et al.; ATS/ERS Task Force for COPD Research. An official American Thoracic Society/European Respiratory Society statement: research questions in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2015;191:e4–e27.
3. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010;11:843–854.
4. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The ensembl variant effect predictor. *Genome Biol* 2016;17:122.
5. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimov A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–249.
6. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 2012;40:W452–W457.
7. Panayiotou C, Solaroli N, Karlsson A. The many isoforms of human adenylate kinases. *Int J Biochem Cell Biol* 2014;49:75–83.
8. da Costa CH, Noronha Filho AJ, Marques E Silva RMF, da Cruz TF, de Oliveira Monteiro V, Pio M, et al. Alpha 1-antitrypsin deficiency in patients with chronic obstructive pulmonary disease patients: is systematic screening necessary? *BMC Res Notes* 2019;12:10.
9. Hersh CP, Dahl M, Ly NP, Berkey CS, Nordestgaard BG, Silverman EK. Chronic obstructive pulmonary disease in α 1-antitrypsin PI MZ heterozygotes: a meta-analysis. *Thorax* 2004;59:843–849.

Copyright © 2019 by the American Thoracic Society



Resistance Heterogeneity and Small Airway Asthma Phenotype



To the Editor:

We read with interest the elegant modeling data of Foy and colleagues (1), who reported a 40% narrowing of small airways was associated with clinically relevant alterations in asthma control and quality of life. Such effects were commensurate with observed responses to biologics on the frequency-dependent heterogeneity of the resistance component of respiratory impedance measured by impulse oscillometry (IOS), where the mean pooled effect on resistance at 5 Hz (R5) – resistance at 20 Hz (R20) was -0.04 (95% confidence interval [CI], -0.03 to -0.05) (kPa/L) · s.

†This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). For commercial usage and reprints, please contact Diane Gern (dgern@thoracic.org).

Originally Published in Press as DOI: 10.1164/rccm.201905-1060LE on July 24, 2019